A coronary heart disease risk model for predicting the effect of potent antiretroviral therapy in HIV-1 infected men

Margaret May,1 Jonathan A C Sterne,1 Martin Shipley,2 Eric Brunner,2 Ralph d’Agostino,3 Peter Whincup,4 Yoav Ben-Shlomo,1 Andrew Carr,3 Bruno Ledergerber,6 Jens D Lundgren,7 Andrew N Phillips,8 Joseph Massaro9 and Matthias Egger1,10*

Accepted 4 June 2007

Background Many HIV-infected patients on highly active antiretroviral therapy (HAART) experience metabolic complications including dyslipidaemia and insulin resistance, which may increase their coronary heart disease (CHD) risk. We developed a prognostic model for CHD tailored to the changes in risk factors observed in patients starting HAART.

Methods Data from five cohort studies (British Regional Heart Study, Caerphilly and Speedwell Studies, Framingham Offspring Study, Whitehall II) on 13 100 men aged 40–70 and 114 443 years of follow up were used. CHD was defined as myocardial infarction or death from CHD. Model fit was assessed using the Akaike Information Criterion; generalizability across cohorts was examined using internal-external cross-validation.

Results A parametric model based on the Gompertz distribution generalized best. Variables included in the model were systolic blood pressure, total cholesterol, high-density lipoprotein cholesterol, triglyceride, glucose, diabetes mellitus, body mass index and smoking status. Compared with patients not on HAART, the estimated CHD hazard ratio (HR) for patients on HAART was 1.46 (95% CI 1.15–1.86) for moderate and 2.48 (95% CI 1.76–3.51) for severe metabolic complications.

Conclusions The change in the risk of CHD in HIV-infected men starting HAART can be estimated based on typical changes in risk factors, assuming that HRs estimated

1 Department of Social Medicine, University of Bristol, UK.
2 Department of Epidemiology and Public Health, University College, London, UK.
3 Department of Mathematics and Statistics, Boston University, USA.
4 Division of Community Health Sciences, St George’s Hospital, University of London, UK.
5 St Vincent’s Hospital and University of New South Wales, Sydney, Australia.
6 Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, Switzerland.
7 Department of Infectious Diseases, Hvidovre University Hospital, Copenhagen, Denmark.
8 Royal Free Centre for HIV Medicine & Department of Primary Care and Population Sciences, University College, London, UK.
9 Department of Biostatistics, Boston University, USA.
10 Institute of Social and Preventive Medicine (ISPM), University of Berne, Switzerland.

* Corresponding author. Institute of Social and Preventive Medicine (ISPM), University of Berne, Finkenhubelweg 11, CH-3012 Bern, Switzerland. E-mail: egger@ispm.unibe.ch
using data from non-infected men are applicable to HIV-infected men. Based on this model the risk of CHD is likely to increase, but increases may often be modest, and could be offset by lifestyle changes.

**Keywords** Highly active antiretroviral therapy, protease inhibitors, non-nucleoside reverse transcriptase inhibitor, coronary heart disease, adverse effects, prognosis, coronary risk factors

**Introduction**

The effectiveness of highly active antiretroviral therapy (HAART) in preventing clinical progression of HIV-1 infection has shifted attention to its adverse effects. HAART is associated with a lipodystrophy syndrome, which is characterized by loss of peripheral subcutaneous fat, accumulation of visceral fat and metabolic complications, including hypercholesterolaemia, hypertriglyceridaemia, insulin resistance, impaired glucose tolerance and type 2 diabetes. These metabolic changes are likely to increase the risk of coronary heart disease (CHD) in patients starting HAART, but this increase has not been well defined to date.

The estimation of changes in CHD risk is hampered by the lack of a suitable comparator population: patients with HIV-1 infection who remain untreated have less advanced disease and differ regarding body mass index (BMI), levels of exercise and smoking. Comparisons with individuals not infected with HIV is also problematic: there are important differences in the prevalence of risk factors such as smoking between HIV-infected and non-infected populations, and HIV-infection itself may affect cardiovascular risk. Examining trends in CHD with time on HAART avoids the possibility of confounding by changes in CHD risk factors associated with increased well-being of patients.

The association of serum lipids and other risk factors with CHD has been studied extensively in populations not known to be infected with HIV. Changes in levels of CHD risk factors associated with HAART are also well documented. In this article we describe the development of a prognostic model, based on five cardiovascular cohort studies of HIV-uninfected men, which is tailored to the changes in cardiovascular risk factors observed in patients starting HAART.

**Materials and methods**

**Cardiovascular cohort studies**

Criteria for selection of cardiovascular cohort studies included the availability of individual person data, a comparable age range, systematic assessment of CHD risk factors and long-term follow up. Data from five cohorts, which are described in detail elsewhere, were included. Briefly, The British Regional Heart Study selected 7735 men aged 40–59 years at random from the age-sex register of one general practice in each of 24 towns in England, Wales and Scotland between January 1978 and June 1980 for a prospective study of cardiovascular disease. The men were followed up for CHD outcomes using two-yearly review of general practice records (including all hospital and clinic correspondence), and additional questionnaires at 5 and 12–14 years after baseline examination, and were flagged with the national mortality register. The Caerphilly Study recruited 2512 men aged 45–59 years between 1979 and 1983 from the town of Caerphilly, South Wales and the adjacent villages. After the baseline examination the men have been seen five times (phase I, II, III, IV and V) over the past 25 years and have been followed up for mortality by flagging with the national death register and for non-fatal CHD outcomes by examining GP and hospital records. The Framingham Offspring Study, based in Massachusetts, USA, included 5124 offspring of 1644 spouse pairs from the original Framingham cohort. At enrolment in 1972, the mean age was 36 years (range 25–70 years) and 52% were female. CHD events were reviewed by a panel of three investigating physicians using all available pertinent records. The Speedwell Study is a prospective cohort study of 2348 men aged between 45 and 63 years of age at first examination who were recruited between 1979 and 1982 from the age-sex registers of 16 general practitioners working from two neighbouring health centres in Bristol, England. The protocol for this study was the same as for the Caerphilly Study. The Whitehall II Study is a cohort of civil servants established between 1985 and 1988 (phase 1). In total, 10 308 civil servants were examined: 6895 men (67%) and 3413 women (33%). CHD was assessed at five phases of data collection using questionnaires, primary care and hospital records. Follow-up for mortality uses flagging with the national death register. The data used in this analysis were all measured at phase 3 (1991–93) and incident CHD events were between phase 3 and phase 5 (1997–99). Blood samples were taken after fasting. Full details of the screening examinations are reported elsewhere.

**Construction of prognostic model**

Data from the five observational cohort studies were used to construct a model for prediction of the risk of CHD. The outcome was CHD events, defined as MI or death from CHD (ICD-9 codes 410-414). Analyses were restricted to men for the follow-up period during which they were aged 40–70 years. Insufficient data were available on women. Men with prevalent CHD, defined as a history of MI, angina (doctor diagnosis, abnormal angiogram, abnormal exercise ECG or Rose Angina questionnaire grade 1 or 2) or ECG abnormalities [Minnesota codes 1 (Q), 4 (ST), 5 (T)] or 7-1 detected at baseline examination, were excluded from the analyses. As baseline measurements might predict less well over time, follow up times were censored at 10 years after their first examination.
The variables considered for inclusion in the model were chosen to reflect known risk factors for CHD such as smoking (never, ex-smoker, pipe or cigar, current cigarette smoker) and particularly those potentially affected by the use of antiretroviral drugs such as systolic blood pressure (BP), total cholesterol, HDL-cholesterol, triglyceride (log transformed to give a normally distributed variable), glucose and diabetes mellitus (modelled as three categories defined by measured glucose with diagnosed diabetics included in the top category: glucose $\leq 5.5$, $5.5 < \leq 7$, $> 7$ mmol/l or a diagnosis of diabetes) and BMI. Baseline hazards were graphed for each cohort separately using a non-parametric approach that estimates the hazard as the first derivative of the smoothed Nelson–Aalen cumulative survival curve. These showed that the baseline hazard of CHD rose exponentially with age and differed between the five cohorts. Therefore we used a parametric model based on the Gompertz distribution (which has an exponentially increasing baseline hazard), with age as the underlying time variable. 19 We also stratified on cohort, to allow for the different rates of CHD in different settings.

Non-linear effects of continuous variables were examined by comparing models with quadratic terms with linear models using the Akaike Information Criterion (AIC). 20 Non-proportionality of hazards was examined using Schoenfeld residuals. We examined all pairwise interactions between covariates using Wald $P$-values and tested for changes in covariate effects across 5-year age bands. Heterogeneity of hazard ratios across the cohorts was examined using the $I^2$ statistic. 21 Candidate prognostic models with and without interaction terms were assessed using AIC as a goodness of fit criterion using first models stratified by cohort and secondly models that pooled the data ignoring heterogeneity across risk factors. Models were then assessed for generalizability across the cohorts using internal–external cross-validation: 22,23 four cohorts are used to fit the prognostic model which is then tested on the left-out cohort, rotating round the omitted cohort. The test statistic is the difference in deviance between the model fitted to the data from the omitted cohort with and without re-estimation of parameters. 23,24 All analyses were performed in Stata (version 9.0, College Station, TX).

Applications of prognostic model
Based on a prospective study of 113 patients (98% men, mean age 40 years) on protease inhibitor (PI) based HAART, 25 we defined typical risk factor profiles for HIV-infected men with moderate or severe metabolic complications. We also generated typical risk factor levels for men of the same age enrolled in the cardiovascular cohort studies. 26 We then estimated hazard ratios comparing HIV-infected, non-smoking men with moderate or severe metabolic abnormalities with an untreated non-smoker and a smoker of the same age.

Results
A total of 13 100 men with 114 443 years of follow up were available for analyses (Table 1). Median [interquartile range (IQR)] age at baseline varied from 47 (43–51) years in the Framingham Offspring study to 55 (51–59) years in the Speedwell study. The proportion of participants with diabetes...
was highest in the Framingham Offspring cohort while the proportion of current smokers was much lower (12%) in the Whitehall II cohort than in the other cohorts. There were 666 CHD events, of which 260 (39%) occurred in participants of the British Regional Heart Study. Rates of CHD events varied from 1.5/1000 years at risk (95% CI 1.2–1.9) at age 40 to 7.5 (95% CI 4.6–12.2) at age 60.

Prognostic model
Table 2 shows crude hazard ratios (HR) for each covariate from univariable models and also the mutually adjusted HR from multivariable models fitted separately on each cohort, together with the P-value for heterogeneity and percentage of total variation due to heterogeneity between studies ($I^2$).

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>British regional heart study</th>
<th>Caerphilly offsprings study</th>
<th>Framingham offspring study</th>
<th>Speedwell heart study</th>
<th>Whitehall II study</th>
<th>Heterogeneity P-value</th>
<th>$I^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Systolic BP (per 20 mmHg)</strong></td>
<td>1.45 (1.30–1.61)</td>
<td>1.39 (1.17–1.66)</td>
<td>1.07 (0.82–1.39)</td>
<td>1.35 (1.20–1.52)</td>
<td>1.57 (1.08–2.28)</td>
<td>0.20</td>
<td>33%</td>
</tr>
<tr>
<td><strong>Total cholesterol (per mmol/l)</strong></td>
<td>1.42 (1.28–1.59)</td>
<td>1.21 (1.06–1.39)</td>
<td>1.56 (1.26–1.92)</td>
<td>1.26 (1.11–1.43)</td>
<td>1.55 (1.25–1.92)</td>
<td>0.10</td>
<td>48%</td>
</tr>
<tr>
<td><strong>HDL-cholesterol (per mmol/l)</strong></td>
<td>0.37 (0.23–0.62)</td>
<td>0.47 (0.27–0.84)</td>
<td>0.38 (0.15–0.99)</td>
<td>0.52 (0.33–0.84)</td>
<td>0.12 (0.05–0.29)</td>
<td>0.004</td>
<td>74%</td>
</tr>
<tr>
<td><strong>Triglyceride (per log mmol/l)</strong></td>
<td>1.37 (1.10–1.70)</td>
<td>1.90 (1.43–2.54)</td>
<td>1.56 (1.06–2.28)</td>
<td>1.67 (1.21–2.29)</td>
<td>3.18 (2.06–4.91)</td>
<td>0.08</td>
<td>53%</td>
</tr>
<tr>
<td><strong>Glucose &gt;5.5 to &lt;7 (mmol/l)</strong></td>
<td>1.22 (0.94–1.58)</td>
<td>1.70 (1.06–2.74)</td>
<td>0.87 (0.48–1.58)</td>
<td>1.44 (0.90–2.31)</td>
<td>1.16 (0.64–2.09)</td>
<td>0.47</td>
<td>0%</td>
</tr>
<tr>
<td><strong>BMI (per kg/m²)</strong></td>
<td>1.18 (0.98–1.42)</td>
<td>1.36 (1.09–1.70)</td>
<td>1.53 (1.06–2.21)</td>
<td>1.18 (0.93–1.50)</td>
<td>2.09 (1.49–2.95)</td>
<td>0.12</td>
<td>45%</td>
</tr>
<tr>
<td><strong>Ex-smoker</strong></td>
<td>0.99 (0.65–1.50)</td>
<td>1.42 (0.70–2.87)</td>
<td>0.59 (0.18–1.94)</td>
<td>1.38 (0.79–2.41)</td>
<td>1.00 (0.57–1.77)</td>
<td>0.58</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Pipe or cigar smoker</strong></td>
<td>1.21 (0.73–1.99)</td>
<td>2.76 (1.33–5.75)</td>
<td>0.97 (0.26–3.62)</td>
<td>1.64 (0.83–3.25)</td>
<td>1.22 (0.29–5.12)</td>
<td>0.61</td>
<td>0%</td>
</tr>
<tr>
<td><strong>Current cigarette smoker</strong></td>
<td>1.86 (1.31–2.66)</td>
<td>2.80 (1.49–5.26)</td>
<td>2.94 (1.15–7.47)</td>
<td>1.77 (1.03–3.06)</td>
<td>1.70 (0.84–3.44)</td>
<td>0.79</td>
<td>0%</td>
</tr>
</tbody>
</table>

**Adjusted HR (95% CI)**

| Systolic BP (per 20 mmHg) | 1.45 (1.29–1.62) | 1.35 (1.13–1.62) | 1.08 (0.80–1.45) | 1.32 (1.17–1.49) | 1.31 (0.88–1.95) | 0.37 | 7% |
| Total cholesterol (per mmol/l) | 1.49 (1.32–1.68) | 1.15 (1.00–1.32) | 1.73 (1.34–2.23) | 1.19 (1.04–1.35) | 1.50 (1.17–1.93) | 0.007 | 72% |
| HDL-cholesterol (per mmol/l) | 0.34 (0.19–0.60) | 0.65 (0.37–1.15) | 0.49 (0.17–1.41) | 0.54 (0.33–0.89) | 0.16 (0.05–0.51) | 0.07 | 55% |
| Triglyceride (per log mmol/l) | 0.72 (0.55–0.95) | 1.37 (0.98–1.94) | 0.79 (0.50–1.27) | 1.14 (0.79–1.66) | 1.05 (0.55–2.02) | 0.07 | 53% |
| Glucose >5.5 to <7 (mmol/l) | 1.16 (0.89–1.52) | 1.55 (0.95–2.53) | 0.63 (0.34–1.19) | 1.18 (0.72–1.91) | 0.95 (0.52–1.73) | 0.15 | 40% |
| Glucose >7 (mmol/l) or diabetes | 1.58 (1.06–2.37) | 1.72 (1.02–3.20) | 0.64 (0.23–1.80) | 1.85 (0.89–3.83) | 2.57 (0.77–8.37) | 0.18 | 36% |
| BMI (per kg/m²) | 0.95 (0.77–1.16) | 1.18 (0.93–1.50) | 1.89 (1.24–2.88) | 0.96 (0.74–1.25) | 1.42 (0.97–2.08) | 0.03 | 52% |
| Ex-smoker | 0.98 (0.65–1.48) | 1.35 (0.67–2.73) | 0.57 (0.17–1.88) | 1.30 (0.74–2.28) | 0.89 (0.50–1.58) | 0.59 | 0% |
| Pipe or cigar smoker | 1.28 (0.77–2.11) | 2.79 (1.34–5.80) | 1.07 (0.29–4.05) | 1.41 (0.71–2.82) | 1.18 (0.28–4.98) | 0.71 | 0% |
| Current cigarette smoker | 1.89 (1.32–2.70) | 3.09 (1.63–5.82) | 3.36 (1.31–8.64) | 1.66 (0.96–2.88) | 1.41 (0.69–2.89) | 0.51 | 0% |

Pooled HRs were derived using fixed-effects and random-effects meta-analysis. In general, HRs from meta-analysis were similar to those from the model using pooled data, with the exception of smoking and diabetes. For smoking and diabetes, the HR from the pooled data were somewhat...
higher than those from the meta-analyses, although
confidence intervals overlapped. No evidence of non-linear
effects was found for the continuous variables systolic BP,
BMI, total cholesterol, HDL-cholesterol and log triglyceride.
There was weak evidence of interaction between systolic
BP and smoking status ($P = 0.15$). The HR for systolic BP
was $1.35$ (95% CI 1.26–1.45) in the model without interactions,
but varied from $1.26$ (1.14–1.39) for current smokers to
$1.51$ (1.33–1.71) for ex-smokers in the model with interactions.
There was some evidence of an interaction between systolic
BP and BMI ($P = 0.02$): increased SBP had a greater effect
in those with BMI $\geq 30$. Tests of the proportional hazards
assumption showed that the HR for current smoker status
($P = 0.002$) decreased with age and therefore we included
an interaction between smoking and age in candidate models.
The effect of being a current smoker was strongest at age
40 and diminished at older ages with nearly all CHD events
in the 40–44 age band occurring in smokers. Triglyceride
was correlated with cholesterol (correlation coefficient $0.36,
P < 0.0005$), and the HR was close to 1 in the multivariable
model.

Models with interaction terms did not fit the data or
generalize across the cohorts substantially better than the
simpler model that did not include interaction terms (models
with interactions AIC 3909 to 3914 compared to AIC 3910
without interaction terms; the deviance difference was well
below the 5% $\chi^2$ reference value for all models indicating no
substantial model misfit when applied to independent data).
We therefore chose the model without interactions as the
final prognostic model. The HR for the pooled data shown in
Table 3 are the HR from this model. Further details of
the prognostic model and the table of model coefficients are
given in the Appendix.

### Application of prognostic model

Table 3 shows the typical risk factor profiles for two HIV-infected men on antiretroviral therapy with moderate or severe
metabolic complications compared with a typical non-infected man from the cardiovascular studies. The CHD HR for the HIV-infected man compared with the control is 1.46 (95% CI 1.15–1.86) for moderate and 2.48 (1.76–3.51) for severe metabolic complications, assuming all are non-smokers. For comparison, the CHD HR for a smoker compared with a non-smoker is 2.04 (1.61–2.57) and therefore the HR for an HIV patient with severe abnormalities who is also a smoker is approximately five compared with the control.

Calculations of the absolute risk of CHD may require
re-calibration of the model for the population in which it is
being used. We followed the method proposed by D’Agostino
et al.$^{27}$ and parameterized the prognostic model by centering
the risk factors on explicitly stated reference values (‘mean
values’$^{27}$) and estimated the constant in the linear predictor
that applies to a fixed absolute 10-year risk of CHD of 3% at
age 40. Examples and details of the reference values and
re-calibration procedure are given in the Appendix.

### Discussion

Thanks to a collaborative effort involving five cardiovascular
cohort studies, a prognostic model that is tailored to the
changes in the cardiovascular risk profile typically observed
in HIV-infected patients initiating HAART could be developed.
Based on the changes in risk factors associated with
starting HAART that have been reported in published studies,
the model suggests that the CHD HR associated with
moderate metabolic abnormalities is about 1.5, less than that
due to smoking (2.0), while the HR associated with severe

#### Table 3

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>Model hazard ratio</th>
<th>Non-smoking HIV-infected men with metabolic complications</th>
<th>Control men</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td>Moderate abnormalities</td>
<td>Severe abnormalities</td>
</tr>
<tr>
<td>Systolic blood pressure (mmHg)</td>
<td>1.34 (1.25–1.44)$^a$</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>BMI (kg/m$^2$)</td>
<td>1.11 (0.98–1.23)$^b$</td>
<td>120</td>
<td>120</td>
</tr>
<tr>
<td><strong>Lipid factors (mmol/l)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total cholesterol</td>
<td>1.31 (1.22–1.40)</td>
<td>5.7</td>
<td>6.2</td>
</tr>
<tr>
<td>HDL-cholesterol</td>
<td>0.46 (0.34–0.61)</td>
<td>0.97</td>
<td>0.85</td>
</tr>
<tr>
<td>Triglycerides</td>
<td>1.00 (0.85–1.18)</td>
<td>3.0</td>
<td>5.0</td>
</tr>
<tr>
<td><strong>Impaired glucose tolerance</strong></td>
<td>1.15 (0.95–1.39)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Diabetes</td>
<td>1.55 (1.16–2.08)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>1.04 (0.81–1.35)</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Pipe or cigar smoker</td>
<td>1.45 (1.06–1.98)</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Current smoker</td>
<td>2.04 (1.61–2.57)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Hazard ratio (HR) for CHD from model</td>
<td>1.46 (1.15–1.86)</td>
<td>2.48 (1.76–3.51)</td>
<td>2.04 (1.61–2.57)</td>
</tr>
</tbody>
</table>

$^a$Hazard ratio for SBP is per 20 mmHg.
$^b$Hazard ratio for BMI is per 5 kg/m$^2$. 

---

abnormalities is about 2.5, only slightly greater than the effect of smoking. A major strength of this simple and generalizable model is the inclusion of variables that are important in the context of HAART-induced metabolic disturbances, BMI and fasting blood glucose. The widely used Framingham or PROCAM risk equations,\textsuperscript{28,29} for example, include the presence or absence of diabetes, but not BMI or blood glucose. The model should be useful to clinicians and patients and help guide decisions on lifestyle changes and preventive drug interventions.

**Applicability of model**

How applicable are our estimates to HIV-infected men? It is well known that the accuracy of prognostic models tends to decline from the data which was used to develop the model to subsequent applications.\textsuperscript{30} We addressed this issue by penalizing model complexity, and by choosing models that generalized best to cohorts omitted from the estimation procedure. Our database included individuals from four community cohorts and one occupational cohort and two countries. The model should therefore, in principle, be transportable to populations other than those from which the cohorts were drawn. Generalizability may also be compromised if important independent predictors are omitted from the model.\textsuperscript{31} We have included classical cardiovascular risk factors but could not include never smokers,\textsuperscript{32} as no information on these was collected in the participating cohort studies. Further, we were not able to include measures of central obesity (e.g. waist–hip ratio), since these were only available in one of the cohorts. Waist–hip ratio is a stronger predictor of CHD risk than BMI and therefore it would be preferable to include waist–hip ratio.\textsuperscript{33} HIV-infected patients may develop coagulation abnormalities such as increased levels of fibrinogen, D-dimer, plasminogen activator inhibitor-1 and tissue-type plasminogen activator antigen, or a deficiency of protein S.\textsuperscript{34} Again, we were unable to consider these factors in our model.

The wisdom, or otherwise, of extrapolating estimates of coronary risk from prospective studies of non-HIV-infected populations to HIV-infected patients with drug-induced metabolic complications remains to be determined.\textsuperscript{35} Interestingly, a collaborative study of more than 23,000 HIV-infected patients from 11 prospective studies recently showed that CHD risk was increased in patients starting HAART including a PI, but not with HAART based on a non-nucleoside reverse transcriptase inhibitor (NNRTI): the relative rate per year of PI exposure was 1.16 (95% CI 1.10–1.23), while for NNRTI based HAART it was 1.05 (0.98–1.13).\textsuperscript{36} We found similar results when applying the prognostic model to men from the Swiss HIV Cohort Study: predicted CHD rate ratios, compared to before HAART was started, were 1.40 (CI 1.13–1.75) and 1.17 (0.95–1.47) for PI- and NNRTI-based HAART, respectively.\textsuperscript{37} A study of 10-year CHD risk estimates obtained from the Framingham risk equation found that HIV-infected patients with fat redistribution had increased risk estimates, but estimated risks were similar when compared to individuals matched for age, sex and BMI from the Framingham Offspring Study.\textsuperscript{38}

**Calibration and re-calibration**

Poor calibration is a universal problem with generalizing prognostic models to populations which differ substantially from those used in fitting the models, either by ethnicity, geography, calendar time, socioeconomic status, diet or any other factor not included in the model which is likely to affect risk of CHD. For example, the Framingham CHD risk functions were based on data from white, middle-class North Americans aged 30–74, with risk factor measurements taken over 30 years ago. These risk functions have been re-calibrated for different ethnic groups,\textsuperscript{27} and for different geographical regions, such as European Mediterranean areas that have lower rates of CHD than the US or Northern Europe.\textsuperscript{39} The method used to re-calibrate the Framingham risk equation, which can also be used for the risk equation presented here, assumes that HRs are constant between cohorts. Whilst some studies have found HRs to be similar across populations,\textsuperscript{29} others, such as the Diverse Populations Collaboration that includes cohorts from North and South America, Scandinavia and Southern Europe, have shown heterogeneity in HRs.\textsuperscript{40}

A simpler method for re-calibration of the Framingham equation was used by Brindle et al.,\textsuperscript{41} who re-scaled the equation for use with British men using data from the British Regional Heart Study. Observed 10 year CHD mortality was compared to the number of events predicted by the Framingham model. The relative degree of overestimation was similar at all levels of CHD risk and so the Framingham equation could be re-scaled by dividing the calculated score for each individual by the amount of over-prediction. The drawback with this method is that it does not take into account differences in prevalence of risk factors in different populations. This is an important factor in re-calibrating risk functions for use with an HIV positive population as their risk of CHD is likely to be higher than the risk for the HIV negative population in the same region, due to higher smoking rates.\textsuperscript{6}

**Relative and absolute changes in CHD risk**

Prognostic models can be used to estimate both the absolute risk of CHD associated with given levels of CHD risk factors, and the CHD HRs associated with given changes in the levels of CHD risk factors. Here, we have focused on the HRs associated with the metabolic abnormalities that have been reported for men on HAART. For patients starting HAART, and their physicians, it will be the absolute risk of CHD that will be of primary importance in considering whether to initiate CHD preventive medication such as statins. Such risks could be estimated using the model presented here, based on patients’ measured risk factor levels, both before and during the first months and years after starting therapy. As illustrated in Figure 1, age will be the most important determinant of CHD risk.

**Modelling of age**

Part of the explanation for variation in the reported effect of CHD risk factors in different populations might be differences in age structure. The modelling of baseline age as a risk factor along with other modifiable risk factors is problematic. CHD risk is extremely low at ages less than 40 and then increases exponentially with age. Therefore we chose to model the effect
of age by including it in the baseline hazard, so that we modelled the change in CHD risk with age, rather than with time in study. This implies that subjects of the same age are compared in assessing the effect of other risk factors, and that it is possible to estimate the cumulative lifetime risk of CHD rather than the risk restricted to the number of years of follow up available in the cohorts. Levels of other risk factors will change with age, and their relative effect on CHD risk might also vary with age. We found some evidence that the effect of smoking was age-dependent, with smoking having a greater effect at younger ages. Finally, we excluded subjects with existing CHD from the data used to fit the prognostic model and from predictions using the model: the proportion of such subjects will vary depending on the age structure of the cohort so that the relevance of our model will decrease in older age groups in whom the prevalence of existing CHD is greater.

Conclusions
In this article we have described a prognostic model for estimating risk of CHD that has been tailored for use with HIV-infected men aged 40–70 years by including risk factors affected by antiretroviral therapy. The estimated increase in CHD HR for patients starting HAART may be directly calculated from the model, using estimated changes in risk factors due to specific antiretroviral regimens. The model also enables the quantification of absolute risk of CHD and, in conjunction with a prognostic model for progression to AIDS or death, potentially allows physicians to assess the benefit and harm of antiretroviral therapy in individual patients. The model requires further development for application to HIV-infected women. In general, the use of the model to estimate absolute risk of CHD will require calibration to specific HIV-infected populations. Calibration and validation of this model and comparison of predictive accuracy with other models, such as the Framingham or PROCAM risk scores, in collaborations of HIV cohorts that have validated cardiovascular endpoints such as the Data Collection on Adverse Events of Anti-HIV Drugs (DAD) Study Group is the subject of future work.

Conflict of interest: None declared.

KEY MESSAGES
- HIV-infected patients on highly active antiretroviral therapy (HAART) often experience metabolic complications, including dyslipidaemia and insulin resistance.
- A model to predict coronary heart disease (CHD) risk in men taking HAART was developed, using data from five cardiovascular cohort studies.
- The estimated CHD hazard ratio was about 1.5 for men with moderate and 2.5 for men with severe metabolic complications.
- Based on this model, increases in CHD risk are modest, and could be offset by lifestyle changes.

References
316 INTERNATIONAL JOURNAL OF EPIDEMIOLOGY


Appendix

Calculation of hazard ratio for CHD from the prognostic model

The Gompertz survival function is parameterized

\[ S(t) = \exp(-\lambda t^\mu) \]

where \( S(t) \) is the probability of survival free of CHD at time \( t \). \( \lambda = \exp(\sum_{i=1}^{n} x_i \beta_i) \) is the linear prediction, constant term always takes the value one.

The CHD HR attributable to treatment with HAART may be calculated from the prognostic model using assumed changes in risk factors \(^{37}\) using the coefficients of the model given in Table A1 below. To calculate the HR substitute the assumed changes in risk factors (in the specified units) in the equation for the linear predictor, \( \lambda \), omitting the intercept terms (i.e. the constant \( X_0 \)) so \( \lambda = \sum_{i=1}^{n} x_i \beta_i \). The HR is then the exponentiated linear predictor \( HR = \exp(\lambda) \). For example, if it is assumed that SBP increases by 5 mmHg, total cholesterol by 2 mmol/l and BMI by 1 kg/m\(^2\), then

\[ \lambda = \left( \frac{0.294 \times 5}{20} \right) + \left( 0.267 \times 2 \right) + \left( \frac{0.101 \times 1}{5} \right) = 0.63 \]

\[ HR = \exp(0.63) = 1.9 \]

95% confidence intervals may be calculated for HRs, but require the variance/covariance matrix (further details available from the authors).

Calculation of absolute risk of a CHD event from the prognostic model

The constants in the model are calibrated for a population with a base 10 year risk of CHD of 3% (or equivalently a 5 year risk of 1.3%) at the reference values of the continuous risk factors and the reference groups of the categorical variables, that is, normal glucose and never smoked. To use the model as it stands—subtract the reference values from the risk factor values of a patient before substituting in the equation of the model. The value of 3% was chosen to be similar to the
Framingham estimate of 10 year risk of hard CHD (myocardial infarction and CHD death) at these values of risk factors and is approximately equivalent to 4% risk of CHD events including angina. In Europe currently used versions of the Framingham equation often include soft CHD endpoints, whereas, in the US the Adult Treatment Panel III (ATP) guidelines on the detection, evaluation and treatment of high blood cholesterol in adults, only hard CHD endpoints are considered. The ATP guidelines (available at http://www.nhlbi.nih.gov/guidelines/cholesterol/index.htm) suggest that hard CHD is between two-thirds and three-quarters of total CHD events.

**Example:** Calculate 5 year risk of CHD for a man aged 50 with SBP = 145, cholesterol = 6, HDL = 1.1, triglyceride = 2, BMI = 27, glucose = 6 and who is a current cigarette smoker.

\[
\Sigma \beta = \left(\frac{145 - 120}{20}\right) \times 0.294 + (6 - 5) \times 0.267 + (1.1 - 1) \times -0.779 + \log(2) \times -0.003 + \left(\frac{27 - 25}{5}\right) \times 0.101 \\
+ 1 \times 0.136 + 1 \times 0.712 + (1) \times -6.148 = -4.705
\]

\[
\lambda = \exp(\Sigma \beta) = \exp(-4.705) = 0.00905
\]

The probability of surviving 5 years free of CHD event at age 50 is calculated as the probability of surviving to age 55 conditional on survival to age 50 (NB subtract the reference age, 40, from these ages). The 5 year risk of CHD is then 5 year risk CHD = 1 – \{S(t = 10) – S(t = 15)\}

\[
S(t) = \exp(-\lambda \gamma^{-1} (\exp(\gamma t) - 1))
\]

Substituting \(t = 10, \gamma = 0.067\) and since risk = 0.05, \(S(10) = 0.95\)

\[
S(10) = \exp(-\lambda \times 0.067^{-1} (\exp(0.067 \times 10) - 1))
\]

\[
0.95 = \exp(-\lambda \times 0.067^{-1} (\exp(0.067 \times 10) - 1)) \log(0.95) = -\lambda \times 0.067^{-1} (\exp(0.067 \times 10) - 1)
\]

\[
\lambda = -0.00720 \times \log(0.95)/(\exp(0.067 \times 10) - 1)
\]

\[
\lambda = 0.0036
\]

We describe an example of recalibration of the model to a population whose reference values are the same as presented in Table A1 other than that mean SBP = 140 mmHg and mean cholesterol = 6.5 mmol/l, and that there are 30% never smokers, 20% ex-smokers and 50% current smokers. The 10 year risk of CHD at these reference values is equal to 5%.

**Step 1:** Calculate the difference in the linear predictor due to the differences between the mean values in the new population and the reference values of the risk factors:

\[
\text{Difference} = \beta_{\text{SBP}} \times (\text{SBP}_{\text{mean}} - \text{SBP}_{\text{ref}})/20 + \beta_{\text{cholesterol}} \times (\text{cholesterol}_{\text{mean}} - \text{cholesterol}_{\text{ref}}) + \beta_{\text{ex-smoker}} \times \text{(proportion ex-smoker)} + \beta_{\text{current-smoker}} \times \text{(proportion current-smoker)}
\]

\[
\text{Difference} = 0.294 \times (140 - 120)/20 + 0.267 \times (6.5 - 5) + 0.044 \times (0.2) + 0.712 \times (0.5)
\]

\[
\text{Difference} = 0.294 + 0.4005 + 0.0088 + 0.356 = 1.0593
\]

**Step 2:** using the known 10 year CHD risk at the mean values of the risk factors in the new population, calculate \(\lambda\) from the following equation

\[
S(t) = \exp(-\lambda \gamma^{-1} (\exp(\gamma t) - 1))
\]

Substituting \(t = 10, \gamma = 0.067\) and since risk = 0.05, \(S(10) = 0.95\)

\[
S(10) = \exp(-\lambda \times 0.067^{-1} (\exp(0.067 \times 10) - 1))
\]

\[
0.95 = \exp(-\lambda \times 0.067^{-1} (\exp(0.067 \times 10) - 1)) \log(0.95) = -\lambda \times 0.067^{-1} (\exp(0.067 \times 10) - 1)
\]

\[
\lambda = -0.00720 \times \log(0.95)/(\exp(0.067 \times 10) - 1)
\]

\[
\lambda = 0.0036
\]

**Step 3:** Calculate the calibration constant in the linear predictor (X8 in Table A1) using the values calculated in steps 1 and 2

\[
\lambda = \exp(\Sigma \beta)
\]

\[
\Sigma \beta = \log \lambda
\]

\[
\text{Difference} + \text{constant} = \log \lambda
\]

\[
\text{Constant} = \log \lambda - \text{Difference} = \log(0.0036)
\]

\[
= -1.0593 = -5.6268 - 1.0593 = -6.6861
\]
The model is now calibrated to the new population. The mean values in the new population should be subtracted from patient risk factor values (instead of subtracting the reference values reported in Table A1) and the new value of the constant is used in calculating the linear predictor. We have assumed that the shape constant $\gamma$ is appropriate to the new population. In principal it would be possible to re-calibrate this parameter too, but this would require estimates of CHD risk at a number of different ages at specified risk factor values in the new population.